Bayesian Vision for Shape Recovery

André Jalobeanu

USRA / RIACS, Bayesian Vision Group (P. Cheeseman)
NASA Ames Research Center MS 269-4
Moffett Field CA 94035-1000, USA
email: ajalobea@riacs.edu

Abstract. We present a new Bayesian vision technique that aims at recovering a shape from two or more noisy observations taken under similar lighting conditions. The shape is parametrized by a piecewise linear height field, textured by a piecewise linear irradiance field, and we assume Gaussian Markovian priors for both shape vertices and irradiance variables. The modeled observation process, equivalent to rendering, is modeled by a non-affine projection (e.g. perspective projection) followed by a convolution with a piecewise linear point spread function, and contamination by additive Gaussian noise. We assume that the observation parameters are calibrated beforehand.

The major novelty of the proposed method consists of marginalizing out the irradiances considered as nuisance parameters, which is achieved by a hierarchy of approximations. This reduces the inference to minimizing an energy that only depends on the shape vertices, and therefore allows an efficient Iterated Conditional Mode (ICM) optimization scheme to be implemented. A Gaussian approximation of the posterior shape density is computed, thus providing estimates of both the geometry and its uncertainty. We illustrate the effectiveness of the new method by shape reconstruction results in a 2D case. A 3D version is currently under development and aims at recovering a surface from multiple images, reconstructing the topography by marginalizing out both albedo and shading.

1. INTRODUCTION

In this work, we investigate the general problem of recovering a shape from a set of corrupted projections (i.e. observations) using an original Bayesian approach. We clearly define the proposed approach, which uses a hierarchy of approximations to marginalize out all the nuisance parameters. We justify the model and algorithm choice (at least experimentally in a 2D case). Future work will extend it to the more realistic 3D framework, the final goal being to recover a 3D surface geometry from multiple 2D images in an efficient and robust way. At first, the lighting and the spatially variable reflectance properties of this surface should be treated as nuisance parameters. Once the surface is estimated this way, we could infer both reflectance properties and lighting.

In 2D, the problem is defined as follows. The surface is a Lambertian emitter defined by a finite curve parametrized by a set of N_v vertices $\mathbf{v} = \{\mathbf{v}^k\}$ (*index* k), and has an irradiance field attached to it, parametrized by a set of irradiance variables $L = \{L^j\}$ (*index* j). There are at least 2 pinhole cameras (parametrized by the set $\mathbf{\Theta} = \{\mathbf{\Theta}^n\}$: position, orientation, etc.) that record the intensity signals $\{X^n\}$ (*index* n) after projecting the irradiance onto a segment, as shown on Fig. 1. The intensity is sampled after convolution with a sampling kernel, or point spread function (PSF), denoted by h. Then it is corrupted by a white additive Gaussian noise of variance σ^2. Our goal is to provide an estimate of the *geometry* $\mathbf{v}$ as well as the *uncertainty* related to this estimate.

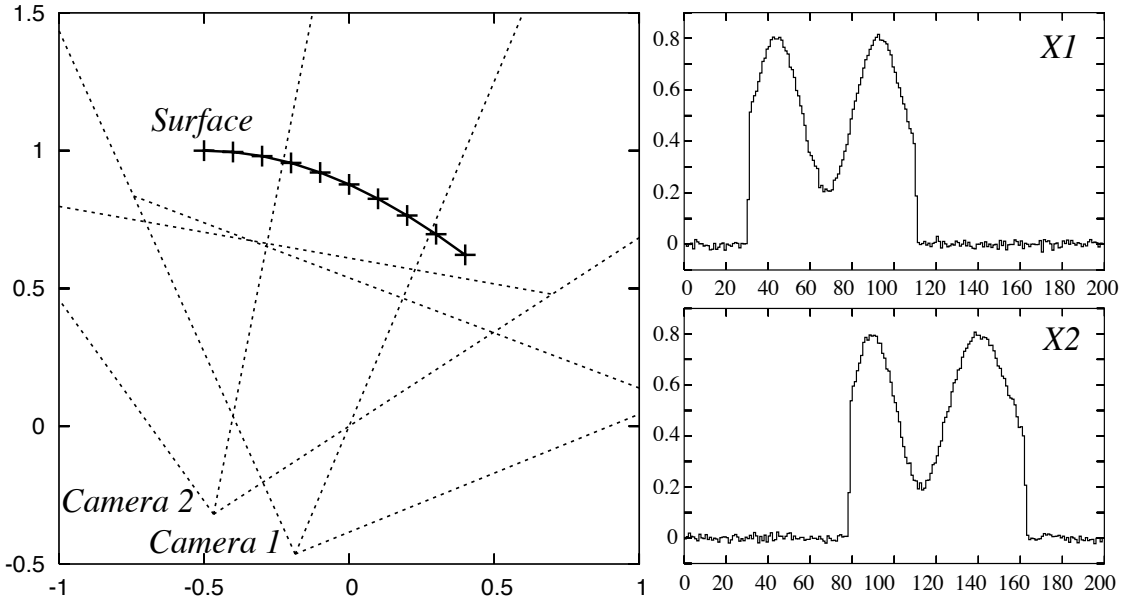

FIGURE 1. Left: the geometry configuration in the object space, showing the 2 cameras. **Right:** the intensity X^n recorded by each camera. The irradiance field $\mathcal{L}$ is not shown here (we choose a smooth sine function with values between 0.2 and 0.8).

To address this problem, we use a Bayesian framework [2]. To simplify it and get an efficient and stable inference procedure, we make a few assumptions about the lighting scheme. We assume similar lighting conditions (as in a stereo setting), which enables us to consider a *single* irradiance field $\mathcal{L}$ attached to the surface, acting as a texture. Usually a surface is parametrized by an albedo field ρ and a reflectance function f, and the irradiance $\mathcal{L} \propto \rho f$. If we assume that we have similar lighting conditions between images, and that f does not depend on the angle between surface and camera, then f is constant from one image to another (though spatially variable), therefore we can re-parametrize by the irradiance.

We can summarize our contributions as follows. First, we derive a model of the unknown surface by choosing an appropriate *parametrization* and efficient *priors* to stabilize the solution. Second, we show how to choose an appropriate *discretization scheme* by understanding the image formation process. Finally, as mentioned above, *irradiances* are used as parameters. We treat them as nuisance parameters and *marginalize* them out, deriving a computationally efficient *hierarchy of approximations*.

2. THE FORWARD PROBLEM

2.1. Generative model and posterior distribution

We assume that all the parameters are random variables governed by a joint probability distribution. The relationships between all these variables are given as a graphical model in Fig. 2 (left), where each arrow represents a conditional density, and each leaf node a prior density. We assume a smoothness prior on $\mathbf{v}$, unless a surface estimate is

given. In all cases we choose to use a Gaussian distribution. The camera parameters $\mathbf{\Theta}$ are assumed known (i.e. Dirac distribution). The observations $\{X^n\}$ are assumed to be independent and corrupted by a zero-mean white Gaussian noise of variance σ^2. Therefore the conditional density of an observation given the surface and camera parameters is an iid Gaussian of mean denoted by I, and variance σ^2.

For each sample point or pixel p, the intensity $I_p(\mathbf{v}, L, \mathbf{\Theta})$ is synthesized from the surface $(\mathbf{v}, L)$ using the camera parameters $\mathbf{\Theta}$ (this is also known as *rendering*, and it is a deterministic process described in Section 2.4).

The likelihood of both surface and camera parameters is:

$$P(\{X^n\}|\mathbf{v}, L, \{\mathbf{\Theta}^n\}) = \frac{1}{Z_\sigma} e^{-U(\mathbf{v},L)} \text{ with } U(\mathbf{v}, L) = \frac{1}{2\sigma^2} \sum_{n,p} \left(I_p(\mathbf{v}, L, \mathbf{\Theta}^n) - X_p^n\right)^2 \tag{1}$$

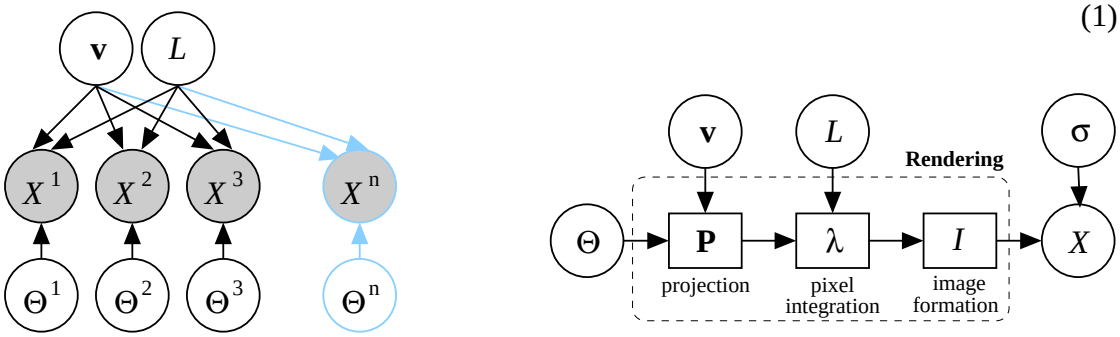

FIGURE 2. The image formation model. **Left:** the multi-image graphical model. **Right:** the detailed model for one image (circles and rectangles respectively represent stochastic and deterministic processes).

2.2. Surface parametrization and topology

We parametrize the geometry by a set of coarse vertices $\{\mathbf{v}^k\}$ using segments as shown on Fig. 3. To constrain the vertices, we parametrize them by a height field $z(x)$: we have $\mathbf{v}^k = (x^k, z^k)$, where the x^k form a fixed grid (uniform in our experiment).

Each segment is subdivided into N_s (equally spaced in our test case) sub-segments containing $N_s - 1$ aligned fine vertices $\{\mathbf{v}^j\}$ which define the irradiance field $\mathcal{L}$. The irradiance has a *higher resolution* than the geometry. There are multiple ways of choosing the irradiance model: it can be piecewise constant between the fine vertices (L^j lies between 2 fine vertices), or piecewise linear (L^j is defined on the vertex $\mathbf{v}^j$), as illustrated by Fig. 3. We will explain in Section 2.4 why a piecewise linear irradiance is preferred.

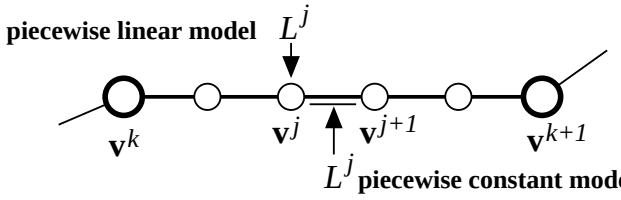

FIGURE 3. The geometry parametrization and topology (subdivided segments), and the 2 possible irradiance models (piecewise linear and piecewise constant).

2.3. The geometry and irradiance priors

To stabilize the solution, in our 2D experiments we use very simple and spatially invariant smoothness priors on both heights and irradiances, corresponding to first order Markov chains (nearest neighbor interactions), with normalizing constants Z_α and Z_β:

$$P(\mathbf{v}) = \frac{1}{Z_\alpha} e^{-\alpha \Phi(\mathbf{v})} \qquad \text{where} \quad \Phi(\mathbf{v}) = \sum_k \left(z^{k+1} - z^k\right)^2 \tag{2}$$

$$P(L) = \frac{1}{Z_\beta} e^{-\alpha \Phi(L)} \qquad \text{where} \quad \Phi(L) = \sum_j \left(L^{j+1} - L^j\right)^2 \tag{3}$$

More complex priors should be used if spatial adaptivity is required, or if discontinuities need to be modeled. Using efficient priors is important when dealing with missing or insufficient data (in real scenarios, some parts of the surface may be hidden) [4].

2.4. The deterministic image formation process

We focus here on the rendering process, or how to obtain discrete pixel intensities I_p from the set of vertices $\mathbf{v}$ and irradiances L in several steps (see Fig. 2 right). Each vertex $\mathbf{v}^j$ is projected according to the following *non-linear* pinhole camera model (this is equivalent to a perspective projection in 3D). Here, $\mathbf{u}$ denotes the observing direction, $\mathbf{T}$ the camera location and F, c are constants related to internal camera parameters:

$$\mathbf{P}_x^j(\mathbf{v}^j, \mathbf{\Theta}) = F \frac{x^j \mathbf{u}_z - z^j \mathbf{u}_x + \mathbf{T}_x}{x^j \mathbf{u}_x + z^j \mathbf{u}_z + \mathbf{T}_z} + c \qquad \text{with} \quad \mathbf{\Theta} = \{\mathbf{u}, \mathbf{T}, F, c\} \tag{4}$$

and $\mathbf{P}^j$ denotes the projected vertex. If we have a linear irradiance model between fine vertices, the projected irradiance field $\mathcal{L}$ is assumed to be also linear (even though the projection is not). The intensity for each pixel I_p is obtained by first convolving $\mathcal{L}$ with a PSF h, then point sampling on a regular grid $\{p\}$: $I_p = (\mathcal{L} \star h)(p)$.

The PSF can be decomposed as a discrete sum of translated sampling kernels Λ, so that I_p is obtained by convolution with a fixed Λ, sampling and then a discrete convolution of the pixel values which can be taken out of the rendering. For simplicity, we ignore this last convolution and assume that $h = \Lambda$.

Since $\mathcal{L}$ is a linear function of $\{L^j\}$, we can write the rendering equation [3] as:

$$I_p = \sum_j \lambda_p^j L^j \qquad \text{where} \quad \lambda_p^j = \text{function of } \mathbf{P}(\mathbf{v}, \mathbf{\Theta}) \tag{5}$$

Then an important question arises: what kernel should we use? The simplest would be a normalized box function $h(x) = 1$ with $x \in [-0.5, 0.5]$. However, we might need a more continuous one, such as the hat function $h(x) = 1 - |x|$ with $x \in [-1, 1]$.

We have implemented the 4 possible combinations of piecewise linear and constant models for both irradiance and sampling kernel, and studied the behavior of the energy $U(\mathbf{v}, L, \mathbf{\Theta})$ relative to one of the images, when one of the vertex heights z^k varies (all others being equal to their true value). Fig. 4 clearly illustrates the fact that choosing

linear models for both irradiance and sampling is the only way to get a smooth energy function, with *continuous derivatives*. For non-boundary vertices, linear irradiance seems to be sufficient; however if we consider vertices located at object boundaries (or occlusion boundaries), the model is discontinuous so we need a continuous sampling scheme, i.e. a continuous PSF h, to achieve the desired smoothness.

Smoothness is important because any deterministic optimization algorithm has little chance of converging to the global optimum if there are discontinuities in the derivatives (on the left figure, the local minima are quite obvious!). For computational reasons, we do not intend to use stochastic techniques such as simulated annealing to get around local minima problems.

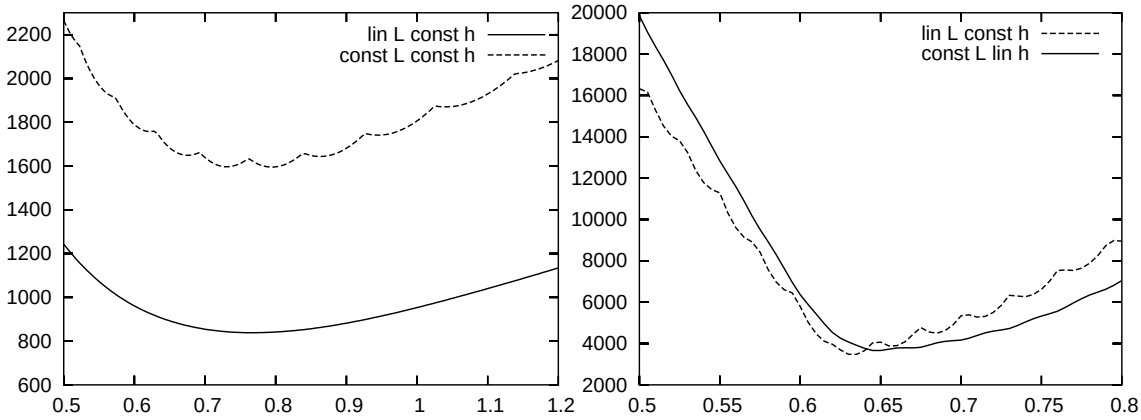

FIGURE 4. Variation of the energy U of Eqn. (1) as a function of one of the vertices (fixed irradiance L). **Left:** non-boundary vertex, piecewise linear vs. piecewise constant irradiance model $\mathcal{L}$. **Right:** boundary vertex, piecewise linear vs. piecewise constant sampling kernel h. Remark: the energy is far more sensitive to boundary than interior vertices.

3. THE INVERSE PROBLEM: BAYESIAN INFERENCE

Computer vision, or model reconstruction from observations, can be seen as the *inverse problem* of rendering. Bayesian inference [2] is an efficient way to deal with such ill-posed inverse problems. In the Bayesian framework, model recovery becomes a parameter estimation problem (more precisely, we estimate parametric pdfs), which is achieved by using existing efficient optimization algorithms.

Using Bayes' rule and the graphical model of Fig. 2 (left) we get the joint posterior:

$$P(\mathbf{v}, L, \{\mathbf{\Theta}^n\}|\{X^n\}) \propto P(\mathbf{v})\, P(L) \prod_i P(X^n|\mathbf{v}, L, \mathbf{\Theta}^n)\, P(\mathbf{\Theta}^n) \tag{6}$$

where the the likelihood is given by Eqn. (1) and the priors by Eqns. (2)- (3).

It is well-known in Bayesian inference that one should integrate the posterior over all unwanted, or nuisance parameters [1], to achieve a good stability: in our case, the camera parameters and the irradiances should be marginalized out. Camera marginalization is simple, i.e. we assumed a Dirac distribution for $\mathbf{\Theta}$. Handling the more realistic case of uncertain camera pose is still work in progress.

3.1. Irradiance marginalization, approximations

We need to calculate the following integral to marginalize out the irradiance L:

$$P(\mathbf{v}|\{X^n\}) = \int_\Omega P(\mathbf{v}, L|\{X^n\})\, dL \propto P(\mathbf{v}) \int_\Omega e^{-U(\mathbf{v},L)-\beta\Phi(L)}\, dL \tag{7}$$

where U is defined in Eqn. (1), $P(\mathbf{v})$ in Eqn. (2) and Φ in Eqn. (3). This can be achieved by using a Laplace approximation, assuming that the integrand can be well approximated by a multivariate Gaussian distribution. The integrand is proportional to the posterior $P(L|\mathbf{v}, \{X^n\})$ (fixed vertices), which is a Gaussian distribution if we assume unbounded irradiances (since I_p is linear in L, U is quadratic in L, and we also have a quadratic penalty Φ defined by Eqn. (3)). For physical reasons, the irradiance is positive and bounded, so the distribution is not rigorously Gaussian. We will assume that the data constrains it to take values far enough from the bounds to ensure the validity of the Laplace approximation. Now we need to calculate the determinant of the covariance matrix $[\Xi]$, and the optimum (MAP) of L given the current geometry $\mathbf{v}$, so that:

$$P(\mathbf{v}|\{X^n\}) \propto P(\mathbf{v})\, \frac{1}{\sqrt{|\Xi|}}\, e^{-U(\mathbf{v},\hat{L})-\beta\Phi(\hat{L})} \tag{8}$$

$$\hat{L}(\mathbf{v}) = \arg\max_L P(L|\mathbf{v}, \{X^n\}) = \arg\min_L \left[U(\mathbf{v}, L) + \beta\Phi(L)\right] \tag{9}$$

$$\left[\Xi^{-1}\right]_{ij} = \frac{\partial^2}{\partial L^i \partial L^j} \left[U(\mathbf{v}, L) + \beta\Phi(L)\right]_{\hat{L}(\mathbf{v})} \tag{10}$$

Experiments have shown that the -log|Ξ| compared to the term $U(\mathbf{v}, \hat{L}) + \beta\Phi(\hat{L})$ has negligible variations when changing vertex heights z^k, so that it can be neglected when optimizing w.r.t. vertices (see Fig. 5 for an illustration).

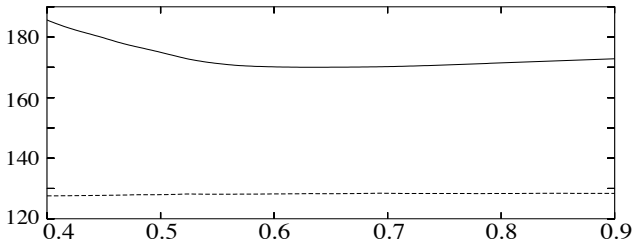

FIGURE 5. Solid line: marginal energy U_m as a function of one of the vertex heights z^k, computed with the proposed approximation; Dashed line: the normalizing constant that was neglected in this computation.

The optimum of Eqn. (9) can be computed by a diagonal Newton-Raphson descent algorithm; a few steps proved to be sufficient for convergence. However, it is preferable to have a *closed-form approximation* of this optimum, which helps calculate the derivatives required for the maximization of the marginal (7), as it will be explained in Section 3.2. We propose one that gives good results in practice, despite the absence of regularization, making optimization search unnecessary. It only involves the weights λ defined by Eqn. (5), and $\{X^n\}$. The accuracy is best when these weights are close to 0 or 1.

$$\hat{L}^j(\mathbf{v}) \simeq \frac{\sum_{n,p} \left(\lambda_p^j\right)^n X_p^n}{\sum_{n,p} \left(\lambda_p^j\right)^n} \tag{11}$$

What makes the marginalization computationally efficient in the proposed approach can be described as a *hierarchy of approximations*: 1) Laplace approximation Eqns. (8)-(10), 2) negligible variations of log-normalizing constant in Eqn. (8), 3) closed-form solution given by Eqn. (11) to the optimal irradiance of Eqn. (9).

Finally we propose the following approximation of the marginalized posterior, so that the entire problem reduces to minimizing the marginal energy denoted by $U_m(\mathbf{v})$:

$$P(\mathbf{v}|\{X^n\}) \propto e^{-U_m(\mathbf{v})} \quad \text{where} \quad U_m(\mathbf{v}) = U\left(\mathbf{v}, \hat{L}(\mathbf{v})\right) + \beta\Phi\left(\hat{L}(\mathbf{v})\right) - \log P(\mathbf{v}) \quad (12)$$

We verify once again that we have made the right choice regarding both irradiance and sampling models: on Fig. 6 the oscillations related to piecewise constant models become quite obvious, and often lead to local minima, and in some cases a bias is noticeable (the global optimum is not the closest to the true value). For boundary vertices the linear sampling scheme is the only one to provide an acceptable marginal energy function.

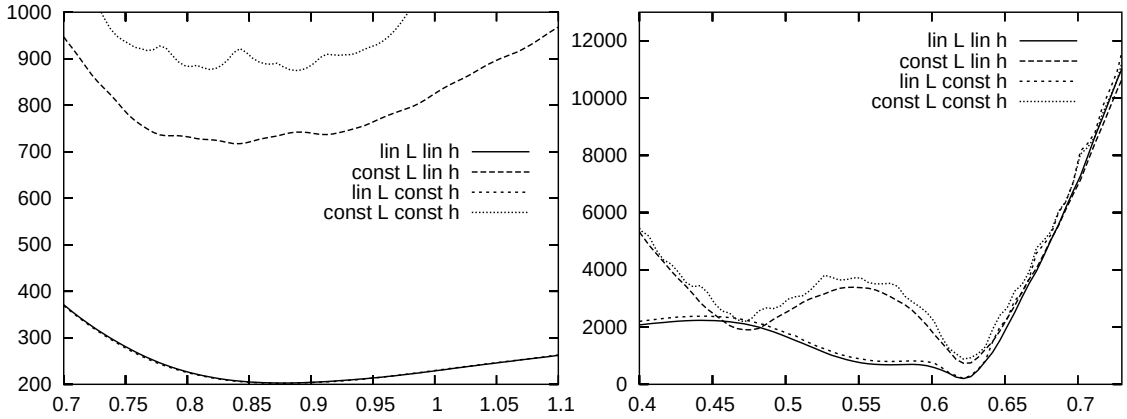

FIGURE 6. Variation of the marginal energy U_m defined in Eqn. (12) as a function of one of the vertices, piecewise linear vs. piecewise constant irradiance $\mathcal{L}$ and sampling h. **Left:** non-boundary vertex. **Right:** boundary vertex, only one curve is smooth on the entire x range (linear $\mathcal{L}$ and h). Remark: the marginal energy is far more sensitive to boundary than interior vertices.

3.2. Computing the derivatives

In order to achieve surface recovery via deterministic minimization of the marginal energy $U_m(\mathbf{v})$, we need to compute the derivatives of this energy w.r.t. vertices. This requires us to compute the derivatives of the rendered intensity I, of the optimal irradiance $\hat{L}$ and of the prior penalty terms Φ. To do this, we use the chain rule, accounting for the graph structure of all variables involved [3].

For the rendered intensity, given the graph structure in Fig. 2 (right) we get the following expression when the irradiance is a function of the vertices:

$$\left[\frac{\partial I_p}{\partial \mathbf{v}^k}\right] = \sum_j \left(\left[\frac{\partial I_p}{\partial \lambda_p^j}\right] \left[\frac{\partial \lambda_p^j}{\partial \mathbf{P}^j}\right] \left[\frac{\partial \mathbf{P}^j}{\partial \mathbf{v}^j}\right] + \left[\frac{\partial I_p}{\partial \hat{L}^j}\right] \left[\frac{\partial \hat{L}^j}{\partial \mathbf{v}^j}\right] \right) \left[\frac{\partial \mathbf{v}^j}{\partial \mathbf{v}^k}\right] \quad (13)$$

This derivative is involved in the first and second derivatives of the marginal energy U_m:

$$\left[\frac{\partial U_m}{\partial \mathbf{v}^k}\right] = \frac{1}{\sigma^2}\sum_{n,p}\left(I_p^n - X_p^n\right)\left[\frac{\partial I_p^n}{\partial \mathbf{v}^k}\right] + \alpha\left[\frac{\partial \Phi(\mathbf{v})}{\partial \mathbf{v}^k}\right] + \beta\sum_j\left[\frac{\partial \Phi(\hat{L})}{\partial \hat{L}^j}\right]\left[\frac{\partial \hat{L}^j}{\partial \mathbf{v}^k}\right] \quad (14)$$

$$\left[\frac{\partial^2 U_m}{\partial \mathbf{v}^k \partial \mathbf{v}^l}\right] \simeq \frac{1}{\sigma^2}\sum_{n,p}\left[\frac{\partial I_p^n}{\partial \mathbf{v}^k}\right]\left[\frac{\partial I_p^n}{\partial \mathbf{v}^l}\right] + \alpha\left[\frac{\partial^2 \Phi(\mathbf{v})}{\partial \mathbf{v}^k \partial \mathbf{v}^l}\right] + \beta\sum_j\left[\frac{\partial^2 \Phi(\hat{L})}{\partial \mathbf{v}^k \partial \hat{L}^j}\right]\left[\frac{\partial \hat{L}^j}{\partial \mathbf{v}^l}\right] \quad (15)$$

where I_p^n refers to the n-th intensity image rendered with the camera parameters $\mathbf{\Theta}^n$. We approximate the second derivatives of U_m by neglecting the contribution of the second derivatives of the intensity and the optimal irradiance $\hat{L}$. Here we can see that we also need to compute the derivative of the optimal irradiance $\hat{L}$, hence the advantage of having a closed-form expression (function of $\mathbf{v}$) such as the one in Eqn. (11).

3.3. The surface recovery algorithm

The goal of the reconstruction algorithm is to provide a Gaussian approximation of the posterior marginal $P(\mathbf{v}|\{X^n\})$. To achieve this, we need to compute the mode $\hat{\mathbf{v}}$ of this distribution, which is equivalent to minimizing the energy U_m defined in Eqn. (12) w.r.t. the geometry $\mathbf{v}$. Then we need a quadratic approximation of this energy around the optimum, hence the second derivatives taken at $\hat{\mathbf{v}}$. The covariance matrix $[\Sigma]$ given in Eqn. (15) provides an measure of the uncertainty on the geometry estimate $\hat{\mathbf{v}}$:

$$\left[\Sigma^{-1}\right]_{kl} = \left[\frac{\partial^2 U_m}{\partial \mathbf{v}^k \partial \mathbf{v}^l}\right]_{\hat{\mathbf{v}}} \quad (16)$$

Keeping the inverse covariance matrix at the end of the reconstruction enables us to build a *recursive update* algorithm, by using this Gaussian approximation as a prior for the next surface estimation process. This way, data can be added to the model *sequentially*. Furthermore, we can remove the restrictive assumptions about the lighting, by processing sets of images with similar lighting conditions simultaneously, and combining the different sets recursively.

The proposed optimization algorithm is iterative, and at each step the rendering process is linearized around the current estimate $\tilde{\mathbf{v}}$ using the intensity derivatives (13):

$$I_p\left(\mathbf{v}, \hat{L}, \mathbf{\Theta}\right) \simeq I_p\left(\tilde{\mathbf{v}}, \hat{L}(\tilde{\mathbf{v}}), \mathbf{\Theta}\right) + \sum_k\left[\frac{\partial I_p}{\partial \mathbf{v}^k}\right]\left(\mathbf{v}^k - \tilde{\mathbf{v}}^k\right) \quad (17)$$

This makes $U(\mathbf{v}, \hat{L})$ a quadratic form in $\mathbf{v}$. Since we choose quadratic penalty functions for the prior defined in Eqn. (2), U_m is also quadratic in $\mathbf{v}$.

Moreover, it has a *first order Markov structure* when there are no occlusions (the dependence is limited to the first order neighbors). Therefore, optimizing this quadratic form is best achieved by an Iterative Conditional Mode (ICM) procedure, which benefits from the local dependence structure. Because of the very limited dependence, using a conjugate gradient in this case is clearly not a good choice, as our experiments have shown. In practice, given the weak dependence between vertices, an *independent optimization* turned out to be very efficient, and was achieved by using a diagonal quasi-Newton descent technique, using the derivatives given by Eqns. (14) and (15):

$$\tilde{\mathbf{v}}^k \leftarrow \tilde{\mathbf{v}}^k - \left[\frac{\partial U_m}{\partial \mathbf{v}^k}\right]\left[\frac{\partial^2 U_m}{\partial (\mathbf{v}^k)^2}\right]^{-1} \tag{18}$$

We also noticed that an accurate optimization of the quadratic form U_m does not help increase the computation speed nor the result quality, since linearization is only a rough approximation. We did not observe any improvement from one to multiple iterations, therefore we do not recommend more than *one step* of descent before recomputing the local quadratic form around the updated geometry estimate.

The proposed algorithm is summarized as follows:

- ***Input:*** $\{X^n\}, \{\mathbf{\Theta}^n\}, \sigma^2$.
- **Initialization:** $\tilde{\mathbf{v}}$ = constant (or a previous estimate if available);
- **Repeat** until convergence:
  - **Irradiance marginalization:** compute $\hat{L}(\tilde{\mathbf{v}})$, Eqn. (11);
  - **Intensity and irradiance derivatives:** Eqn. (13), derivative of Eqn. (11);
  - **First and second derivatives of U_m:** Eqns. (14)-(15);
  - **One step diagonal quasi-Newton update:** Eqn. (18).
- **Inverse covariance estimation:** Eqn. (16).
- ***Output:*** geometry $\hat{\mathbf{v}} = \tilde{\mathbf{v}}$, inverse covariance $[\Sigma^{-1}]$.

The Bayesian approach enables us to automatically estimate the regularization parameters α and β, as well as the subdivision level N_s (i.e. model selection, via Bayes factors). This has been investigated in our experiments, but it is not described in this paper. Moreover, the segment subdivision level could be selected locally and dynamically (in the inner loop of the algorithm), depending on the current local geometry.

4. 2D EXPERIMENT: RESULTS AND CONCLUSIONS

The reconstruction algorithm has been successfully applied to the problem described in the introduction, and the results are shown on Fig. 7. We have $\sigma = 0.01$ in this experiment (1% of the intensity range, since $I_p \in [0, 1]$). The reader can evaluate the reconstruction quality by observing the estimated vertices and the error bars (corresponding to a marginal posterior probability greater than 0.1, computed using the estimated inverse covariance). The plot on the right shows the joint distribution of 3 pairs of vertices conditioned upon all others: only nearest neighbors interact and can therefore be strongly correlated (the correlation decays exponentially with the distance). The coarse segments are larger than pixels, otherwise there would be longer range interactions. This means that *the inverse covariance matrix is sparse*: it is possible to store it along with the surface estimate and to use it as a prior for subsequent inferences.

The main conclusions of the 2D experiments are: 1) *continuous models* for both irradiance and sampling should be used to ensure the continuity of the energy functions; 2) *irradiance marginalization* reduces the problem dimension from $(N_s + 1)N_v$ irradiances+vertices to N_v vertices, also strongly reducing the *interaction structure*; 3) we observed that marginalization also makes the energy landscape more *quadratic*, making the use of Newton-like techniques appropriate.

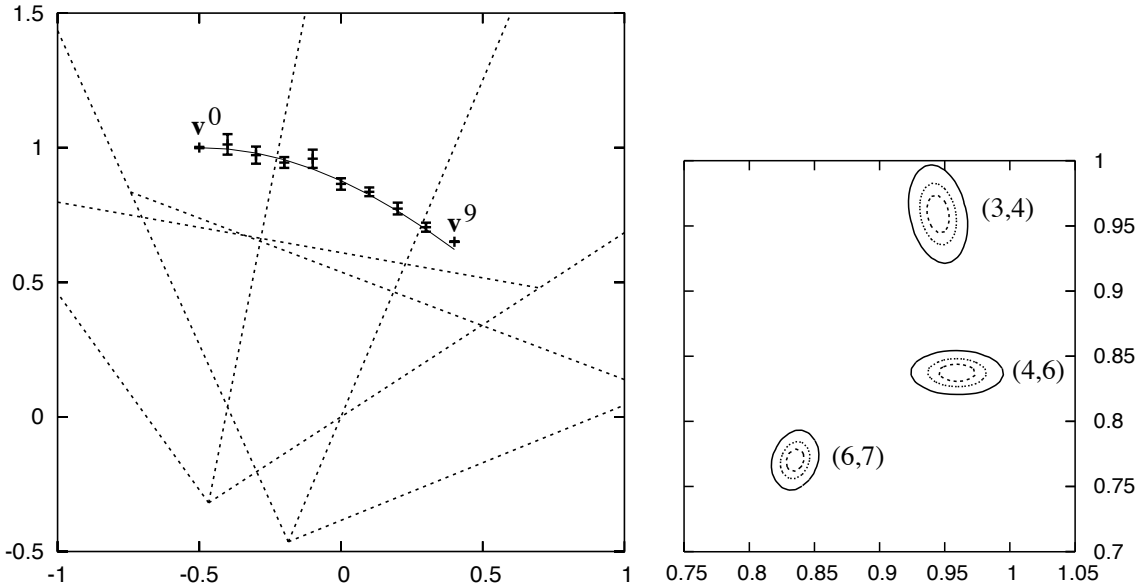

FIGURE 7. The inferred vertex heights (10 iterations, parallel quasi-Newton optimization). **Left:** error bars = inferred geometry (probability > 0.1), solid line = ground truth, dashed lines = the two camera configurations. **Right:** joint conditional posterior distribution of 3 vertex heights given all others (the indices are shown), illustrating the pairwise interactions.

5. EXTENSION TO 3D AND FUTURE WORK

To extend this promising approach to a more useful 3D framework, we use subdivided triangular meshes. *Hidden surface removal* is a crucial to rendering in 3D (we assumed there were no occlusions in our 2D experiments). This is achieved through a recursive approach to subtract triangles from a polygon [4]. Rendering using continuous irradiance and PSF is made possible by computing the moments of the visibility polygons [3].

In 3D, the goal is to first infer the scene geometry, using a 3D equivalent of the method described in this paper (see [3]), then to infer the albedo and reflectance using the estimated geometry (empirical Bayes approach). Simultaneous reconstruction and *camera calibration* could also be addressed through marginalization. To achieve accurate surface recovery, designing and studying realistic *prior models* [4] is needed. Bayesian *model selection* should be considered if a spatially adaptive mesh subdivision is required.

The approach presented here should be sufficiently general to apply to other inverse problems related to shape recovery from projections (e.g. 2D or 3D tomography).

REFERENCES

1. J.O. Berger, B. Liseo, and R. Wolpert. Integrated likelihood methods for eliminating nuisance parameters. *Statistical Science*, 14(1):1–28, 1999.
2. Jose M. Bernardo and Adrian F. M. Smith. *Bayesian Theory*. John Wiley & Sons, Inc., 1994.
3. A. Jalobeanu. Mathematics of accurate super-resolution rendering. *RIACS Research Report*, in preparation 2004.
4. A. Jalobeanu, F.O. Kuehnel, and J.C. Stutz. Modeling Images of Natural 3D Surfaces: Overview and Potential Applications. In *GMBV'04 (part of CVPR'04)*, Washington, DC, July 2004.